

CycleGAN と顕著性マップを用いた写真の鉛筆画による抽象化表現

小林 剛豊†

茅 暁陽‡

山梨大学工学部コンピュータ理工学科†

山梨大学大学院総合研究部‡

1 はじめに

人間が鉛筆画を描く場合は風景の重要な（注意をひく）部分を詳細に、他の部分を省略して描くという抽象化が行われる。本研究の目的は人間の手を介さず写真から抽象化された鉛筆画を生成することである。そのため、近年注目されている Generative Adversarial Networks (GAN) [3] による画像変換技術と顕著性マップ生成技術を組み合わせ、入力画像から抽象化された鉛筆画を生成する技術を提案する。

2 関連研究

LIC 法を利用した鉛筆画生成法

LIC 法とはテキストチャベースの流れ場の可視化手法である。写真からベクトル場を生成し、ホワイトノイズ画像をそのベクトル場に沿ってぼかすことで流れの軌跡を可視化した画像が得られる。茅ら[1]はこの軌跡が鉛筆画のストロークに似ていることを利用して LIC 法を鉛筆画生成に応用した。この生成法では画像の重要な部分が強調された鉛筆画を生成することができないという課題があった。そこで島ら[2]は画像中の色、輝度、方向性の周辺差分(特徴)から誘目度を予測する顕著性マップと LIC 法を用いた鉛筆画生成法を組み合わせることで重要な部分が強調された鉛筆画を生成することを可能にした。しかし LIC 法を用いた生成法ではベクトル場とホワイトノイズ画像を畳み込むことで鉛筆画のストロークに近似させているだけであるため、生成画像は実際の鉛筆画と比べてリアルさに欠ける。この問題を踏まえて本研究では GAN により鉛筆画を教師データとして学習を行い、よりリアルな鉛筆画を生成する。また島らが使用した顕著性マップは低レベルの特徴から計算されるため意味論的な情報を考慮した予測ができない。そこで本研究では人間の視線データを学習に用いて、GAN で生成した顕著性マップを使用することで、人間の知覚をより正確に反映した誘目度の予測を行う。

3 GAN による画像変換技術

GAN[3]とは

教師無し学習モデルの一種であり、ランダムノイズを入力として訓練データに近い画像を生成する生成器とその出力と訓練データを区別する判別器を敵対させ学習を行うことで生成器にランダムノイズから訓練データに近い画像を生成する方法を獲得させる。オリジナルの GAN モデルは生成される画像の属性を制御することができなかったが Mirza らが Conditional GAN (CGAN) [4] を提案し、生成器と判別器の入力に条件ベクトルを加えることで、条件ベクトルに見合う画像を生成可能にした。CGAN の条件ベクトルを画像としてその画像から別の画像を生成する手法を Pix2Pix [5] も開発された。訓練データに絵画を用いれば、Pix2Pix を用いて画像を絵画調画像に変換することが可能となる。

Pix2Pix で画風変換を行うには訓練データとして変換前の画像と変換後の画像のペアが必要である。しかし、通常そのようなペアを入手することが困難である。この問題を解決するため Zhu らが CycleGAN [6] という手法を提案した。CycleGAN の構造を図 1 に示す。GAN では生成器と判別器 1 組を使い学習を行うのに対して、CycleGAN では $X \rightarrow Y$ への変換を行う生成器 G と $Y \rightarrow X$ への変換を行う生成器 F 、それぞれの判定を行う判別器 Y と判別器 X を用いて学習を行う。訓練データがペアとなっていない場合、判別器の出力から計算される誤差だけでは正しいマッピングの方法を学習することはできない。この問題を解消するために CycleGAN では図 2 に示す一貫性誤差を用いる。図 2(a) では x を生成器 G, F に通し、 $X \rightarrow Y \rightarrow X$ の変換を行っている。その結果と x との絶対値の差分を計算しこれを x についての循環誤差とする。また図 2(b) においても x と同様に y についての循環誤差を計算する。この 2 つの循環誤差の合計が一貫性誤差である。学習が進むとこの誤差が小さくなり、生成器は入力を考慮して画像を生成するようになる。

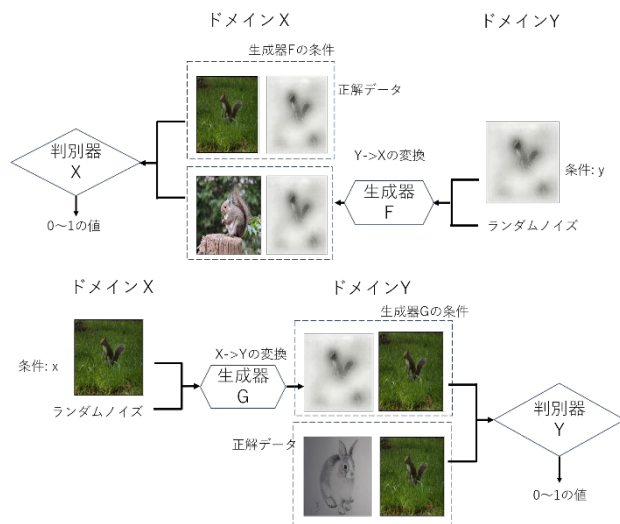


図 1 CycleGAN モデル

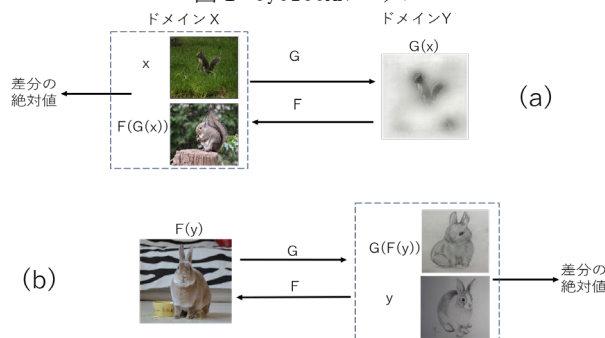


図 2 一貫性誤差

「Abstracted Pencil Image Generation using CycleGAN and Saliency Map」

† Taketo Kobayashi, Department of Computer Science and Engineering, University of Yamanashi

‡ Xiaoyang Mao, Graduate School, University of Yamanashi

4 顕著性マップによる抽象化表現

CycleGAN のみでは抽象化された鉛筆画を生成することはできない。そこで本研究では画像中の注視領域を予測する顕著性マップと組み合わせることでこの問題を解消する。顕著性とは注意を引く度合いを意味し、顕著性マップとは画像内の各領域の顕著性を視覚化したものである。画像中で視線が集まる位置では高い値を示し、視線の集まり度合いを表現する。

SalGAN

従来の顕著性マップ生成法では画像内の低レベルの特徴を用いるため、意味論的な情報による注意への影響をモデリングできないという問題があった。近年視線追跡データを訓練データとする深層学習モデルによる顕著性マップ生成が注目を浴びている。本研究では GAN を利用した顕著性マップ生成モデル SalGAN[7]を用いる。

5 提案手法

図3に提案する鉛筆画生成モデルを示す。

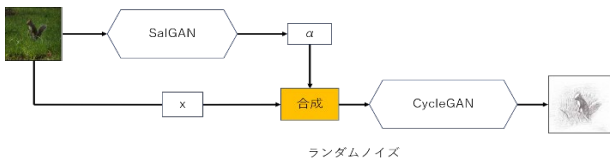


図3 提案手法

SalGAN から得られる顕著性マップと入力画像を合成した後、CycleGAN に入力することで顕著な領域の情報のみを残した抽象化表現を得る。抽象化された鉛筆画においては顕著でない領域はストロークがあまり描かれないため、白に近づけさせる必要がある。一方顕著な領域であるほど画像の特徴を強調する必要がある。このような効果を実現するために提案手法では以下の式を用いて顕著性マップと入力画像の合成を行う。

$$W = (255 - \alpha^\gamma) + \alpha^\gamma * x \quad (1)$$

ここで x は入力画像のピクセル値、 α は顕著性マップのピクセル値である。 γ は強調の度合いを制御するパラメータである。 γ を小さくすることで顕著な領域の周辺も描かれることになる。実験では CycleGAN をアンペアの写真と鉛筆画 150 組を用いて学習させた。SalGAN については SALICON challenge 2017[8]のデータセット(10000 組)で事前に学習されたモデルを使用した。提案手法で写真を鉛筆画に変換した結果を図4に示す。

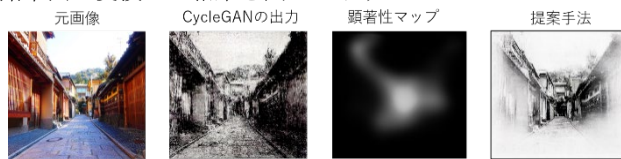


図4 生成結果

図4から提案手法による生成画像では顕著な領域が詳細に、それ以外が省略されるように描かれていることがわかる。次に γ の値を変化させた生成画像の比較を以下の図5に示す。

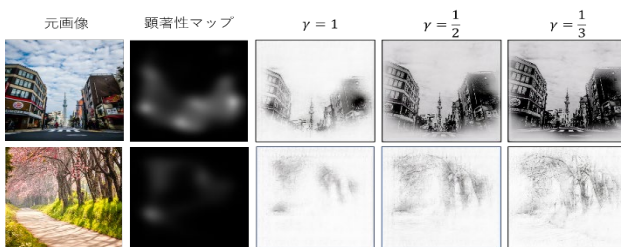


図5 γ の効果

比較に用いた 2 種類の画像には顕著性マップにおいて大きな違いがみられる。上段の画像は顕著な領域が広範囲に分布しかつ顕著性が強く、下段の画像では顕著な領域が狭く顕著性が弱い。図5において γ が 1/3 の列に注目すると下の元画像からは適切に抽象化された鉛筆画が生成されているが、上の元画像から生成された画像では抽象化の効果が薄れてしまい画面全体を描いてしまっていることがわかる。提案手法では γ を小さい値に設定することで顕著性マップの非線形マップによる抽象化を和らげ、より多くの領域を詳細に描くことができる。これは顕著な領域が狭い画像や顕著性が弱い画像に対して有効な手段である。一方顕著性マップにおいて顕著な領域が広範囲である画像に対しては、 γ を大きめの値に設定することで、抽象化の度合いを高めることができる。最後に提案手法と畠らの手法を比較した結果を以下の図6に示す。



図6 既存手法との比較

畠らの手法では水面に映る建物がぼやけて描かれているが、提案手法では γ の値を調整することで詳細に描画する領域を広げることにより水面に映る建物もはっきり描かれていることがわかる。

5 おわりに

今回は CycleGAN と SalGAN を利用し、顕著度によって描画の詳細度を制御する手法を提案した。課題として式(1)による合成方法では、顕著でない領域に雲などの白い物体が含まれると、雲の周辺の領域が白に近づくため雲がほとんど見えなくなってしまうという問題が挙げられる。今後の研究ではコントラスト強調などを用いて白い物体と周辺領域の境界を強調することでこの問題を解消していかうと考えている。

7 参考文献

- [1] 茅 暁陽, 長坂 好恭, 山本 茂文, 今宮 淳美, “LIC 法を利用した鉛筆画の自動生成法”, 芸術科学会論文誌
- [2] 畠 康高, 豊浦 正広, 茅 暁陽, “Saliency Map と LIC 法を利用したアクセントのある鉛筆画生成法”, 電子情報通信学会論文誌
- [3] Goodfellow Ian J., Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio “Generative Adversarial Networks”, NIPS 2014
- [4] Mehdi Mirza, Simon Osindero, “Conditional Generative Adversarial Nets”, arXiv preprint arXiv:1411.1784, 2014
- [5] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, Alexei A. Efros, “Image-to-Image Translation with Conditional Adversarial Networks”, CVPR 2017
- [6] Jun-Yan Zhu, “Unpaired image-to-image translation using cycle-consistent adversarial networks”, ICCV2017
- [7] Junting Pan, Cristian Canton-Ferrer, Kevin McGuinness, Noel E. O’Connor, Jordi, Elisa sayrol, Xavier Giro-i-Neito, “SalGAN: visual saliency prediction with adversarial networks”, CVPR 2017
- [8] <http://salicon.net/challenge-2017>